

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

日 本 国 特 許 庁
PATENT OFFICE
JAPANESE GOVERNMENT



別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日

Date of Application:

2000年 3月 9日

出 願 番 号

Application Number:

特願2000-065485

出 願 人

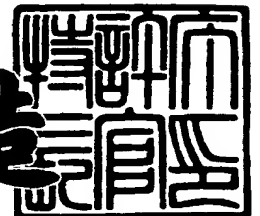
Applicant (s):

日本電気株式会社

2000年12月 1日

特許庁長官
Commissioner,
Patent Office

及 川 耕 造



出証番号 出証特2000-3098031

【書類名】 特許願

【整理番号】 66206163

【提出日】 平成12年 3月 9日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 11/20
G06F 15/16

【発明者】

 【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

 【氏名】 鈴木 勉幸

【特許出願人】

 【識別番号】 000004237

 【氏名又は名称】 日本電気株式会社

【代理人】

 【識別番号】 100080816

 【弁理士】

 【氏名又は名称】 加藤 朝道

 【電話番号】 045-476-1131

【手数料の表示】

 【予納台帳番号】 030362

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

 【物件名】 要約書 1

 【包括委任状番号】 9304371

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 冗長構成クロスバスイッチシステム

【特許請求の範囲】

【請求項 1】

複数のノードに対してノード間の接続に必要とされる複数のクロスバスイッチに、さらに少なくとも1つのクロスバスイッチを余分に有する冗長構成のクロスバスイッチを備え、

前記冗長構成のクロスバスイッチのうち、第1のクロスバスイッチは、その入力端に、前記複数のノードの各ノードの複数（N個）の出力のうち1番目の出力を入力し、前記余分に設けられた1つのクロスバスイッチは、その入力端に、前記複数のノードのN個の出力のうちN番目の出力を入力し、

残りのクロスバスイッチの各々は、前記複数のノードの各ノードから、前記ノードのN個の出力のうち、前記各クロスバスイッチに対応する順番の隣り合う2つの出力をそれぞれ入力とするM個の選択回路をそれぞれ備え、前記M個の選択回路の出力を入力端に入力し、

前記複数のノードの各ノードの入力端には、前記冗長構成のクロスバスイッチの互いに隣り合う2つのクロスバスイッチの出力のうち前記各ノードに対応する順番の2つの出力をそれぞれ入力とするN個の選択回路を備え、

前記各選択回路は、クロスバスイッチの障害処理を行う障害処理回路から出力される選択制御信号により2入力的一方を選択出力し、一のクロスバスイッチに障害発生時、該障害クロスバスイッチを未使用とする、ことを特徴とする冗長構成クロスバスイッチシステム。

【請求項 2】

第1乃至第M（ただし、Mは2以上の所定の整数）のノードに対してノード間の接続に必要とされるN個のクロスバスイッチにさらに1つのクロスバスイッチを備えたN+1の冗長構成のクロスバスイッチを備え、

第1のクロスバスイッチは、M個の入力端に、前記第1乃至第MのノードのM個の出力のうち第1の出力を入力し、

第N+1のクロスバスイッチは、M個の入力端に、前記第1乃至第Mのノード

のN個の出力のうち第Nの出力を入力し、

前記第I（ただし、Iは2以上且つN以下の整数）のクロスバスイッチは、M個の入力端に、前記第1乃至第Mのノードの出力のうち、前記第Iのクロスバスイッチに対応する順番の第I-1の出力と第Iの出力をそれぞれ入力とするM個の選択回路を備え、

前記第Jのノード（ただし、Jは1以上且つM以下の整数）の入力端には、第1乃至第N+1のクロスバスイッチのうち隣り合う2つのクロスバスイッチのJ番目の出力をそれぞれ入力とするN個の選択回路を備え、

前記各選択回路は、クロスバスイッチの障害処理を行う障害処理回路から出力される選択制御信号により、2入力的一方を選択出力し、一のクロスバスイッチに障害発生時、該障害クロスバスイッチを未使用とする、ことを特徴とする冗長構成クロスバスイッチシステム。

【請求項3】

第1乃至第Mのノード（ただし、Mは2以上の所定の整数）に対してノード間の接続に必要とされるN個クロスバスイッチにさらに1つのクロスバスイッチを備えた、第1乃至第N+1のクロスバスイッチを備え、

前記第1乃至第Mのノードの各ノードは、出力端から第1乃至第Mの出力信号を出力するとともに、入力端から、第1乃至第Nの入力信号を入力し、

第1のクロスバスイッチは、M個の入力端に、前記第1乃至第Mのノードの各ノードの1番目の出力信号を入力し、

第N+1のクロスバスイッチは、M個の入力端に、前記第1乃至第Mのノードの各ノードのN番目の出力信号を入力し、

第I（ただし、Iは2以上且つN以下の整数）のクロスバスイッチは、M個の入力端のそれぞれに対して、前記第1乃至第MのノードのI-1番目の出力信号とI番目の出力信号の2つの信号をそれぞれ入力とするM個の選択回路を備え、

第Jのノード（ただし、Jは1以上且つM以下の整数）のN個の入力端には、第1乃至第N+1のクロスバスイッチのうち隣り合うクロスバスイッチである、K番目とK+1番目（ただし、Kは1以上且つN以下の整数）のクロスバスイッチのJ番目の出力ポートの出力をそれぞれ入力とするN個の選択回路を備え、

前記各選択回路は、クロスバスイッチの障害処理を行う障害処理回路から出力される選択制御信号により、2つの信号の一方を選択出力し、一のクロスバスイッチに障害発生時、該障害クロスバスイッチを未使用とする、ことを特徴とする冗長構成クロスバスイッチシステム。

【請求項4】

計算機システム内のCPUとメモリの間を接続するか、あるいは複数のノードで構成される計算機システムにおける各ノード間を接続するクロスバスイッチを備えたシステムにおいて、

システムにN個必要なクロスバスイッチを冗長分1個を含めてN+1個備え、

システム障害発生時に、障害処理回路がクロスバスイッチの障害を認識した場合、システムの再立ち上げ後は、クロスバスイッチの入出力部分に設けた選択回路を、前記障害処理回路が制御することで、障害の発生したクロスバスイッチを未使用とし、冗長分のクロスバスイッチを使用するように制御する手段を備えたことを特徴とする冗長構成クロスバスイッチシステム。

【請求項5】

前記各ノードは、バイト単位にNバイトのデータを入出力する、ことを特徴とする請求項3記載の冗長構成クロスバスイッチシステム。

【請求項6】

前記障害処理回路が、前記第1乃至第N+1のクロスバスイッチについて、障害の有無を保持するN+1ビットのクロスバ障害情報保持レジスタと、前記クロスバ障害情報保持レジスタの値に基づき、前記各選択回路に対して選択制御信号を出力する選択回路制御出力回路と、複数のクロスバスイッチの障害発生時、多重障害発生をシステム制御部に通知する多重障害検出部と、を備えていることを特徴とする請求項1乃至4のいずれか一に記載の冗長構成クロスバスイッチシステム。

【請求項7】

前記クロスバスイッチが、計算機内のCPUとメモリ間の接続、もしくは、CPUとメモリを備え前記クロスバスイッチを介してリモートノードのメモリへのアクセスを行うマルチノードシステムのノード間のスイッチングを行う、ことを

特徴とする請求項 1 乃至 6 のいずれかーに記載の冗長構成クロスバスイッチシステム。

【請求項 8】

前記 M と前記 N とが等しい値とされていることを特徴とする請求項 1、2、3 のいずれかーに記載の冗長構成クロスバスイッチシステム。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、計算機システムに関し、特に、計算機の CPU とメモリの間、あるいはマルチノードで構成される計算機システムにおける各ノード間を接続するクロスバスイッチを冗長構成としたシステムに関する。

【0 0 0 2】

【従来の技術】

近時、計算機システムにおいて、クロスバスイッチは、例えば計算機の CPU とメモリの間、あるいはマルチノードで構成される計算機システムにおける各ノード間をスイッチング接続するために用いられている。クロスバスイッチは、ビットスライスやバイトスライス等に対応して、複数の L S I、または、カード、あるいは複数のクロスバスイッチブロックを含む 1 チップ L S I 上等に構成される。

【0 0 0 3】

【発明が解決しようとする課題】

しかしながら、クロスバスイッチに障害が発生した場合、クロスバスイッチは、CPU やメモリ等のように、デグレードすることができない。すなわち、障害 CPU 等は、システム再立ち上げ時等にシステムから切り離し、残りの CPU で処理が行われるが、障害クロスバスイッチをデグレードした場合、計算機の CPU とメモリの間、あるいはノード間を接続することができず、その結果、システムは動作しないことになる。

【0 0 0 4】

そこで、クロスバスイッチの障害部分等を交換するまでの間、システムの復旧

を行うことができず、システムダウン時間が長くなるという問題点を有している。

【 0 0 0 5 】

また、この問題点を回避するために、クロスバススイッチの部分で完全二重化する構成としたシステムも存在するが、高価なシステムとなり、実用に供し得ない。

【 0 0 0 6 】

なお冗長構成のクロスバススイッチとして、例えば特開平 7 - 2 6 4 1 9 8 号公報には、 $N \times N$ のクロスバススイッチ装置において、 $(N + 1) \times N$ 構成のクロスバススイッチとして、 $N + 1$ 番目の予備入力方路を設け、 N 本の現用方路のうち一本の異常が検出された場合、空間スイッチにより予備方路に切り替えてデータを迂回させるようにした構成が開示されている。また例えば特開平 1 1 - 3 3 1 3 7 4 号公報には、ATM 交換機等の用いられるクロスバススイッチとして、上記特開平 7 - 2 6 4 1 9 8 号公報に記載された冗長構成のクロスバススイッチの問題点（装置運用状態のまま異常箇所を復旧処置できない）を解消するための構成として、 N 個のポートを収容し、任意の 2 ポート間スイッチ機能を実現するクロスバススイッチ部が、複数枚の $N \times N$ クロスバススイッチカードからなり、クロスバススイッチ部と複数の現用方路及び少なくとも一つの予備方路に接続されていた N 個のポート部と、スイッチ信号に応答して設定されたクロスバススイッチ部のスイッチの異常を検出してスイッチ異常検出信号を出力する接続制御部とを備え、ポート部がスイッチ異常検出信号に応答して少なくとも 1 つの現用方路を少なくとも 1 つを予備方路に切り替える構成が開示されている。

【 0 0 0 7 】

上記特開平 1 1 - 3 3 1 3 7 4 号公報に記載されたクロスバススイッチ装置においては、異常が検出された $N \times N$ クロスバススイッチカードに伝送されるパケットデータを、予備方路に迂回させ、予備クロスバススイッチカードへの切り替えを行うものであり、クロスバススイッチカードに障害発生時、パケットデータの迂回先（切り替え先）は、予め定められた予備クロスバススイッチカードとされている。このため、元のクロスバススイッチカードと、迂回先の予備クロスバススイッチカー

ドの配置如何によっては、データ転送のための遅延時間が問題となり、特に、動作周波数の高速化が著しい計算機において、このデータ転送の遅延時間の変動は重大な問題となる。

【0008】

したがって本発明は、上記問題点に鑑みてなされたものであって、その目的は、クロスバススイッチに障害が発生した場合の早急なシステムの復旧を、比較的安価なシステムで実現する、システムを提供することにある。これ以外の本発明の目的、特徴、利点等は、以下の実施の形態の記載から、当業者には、直ちに明らかとされるであろう。

【0009】

【課題を解決するための手段】

前記目的を達成する本発明は、システムにN個必要なクロスバススイッチを冗長分1個を含めてN+1個備え、システム障害発生時に、障害処理回路がクロスバススイッチの障害を認識した場合、システムの再立ち上げ後は、クロスバススイッチの入出力部分に設けた選択回路を、前記障害処理回路が制御することで、障害の発生したクロスバススイッチを未使用とし、冗長分のクロスバススイッチを使用するように制御する手段を備えたものである。

【0010】

より詳細には、本発明は、第1乃至第M（ただし、Mは2以上の所定の整数）のノードに対してノード間の接続に必要とされるN個のクロスバススイッチにさらに1つのクロスバススイッチを備えたN+1の冗長構成のクロスバススイッチを備え、第1のクロスバススイッチは、M個の入力端に、前記第1乃至第MのノードのN個の出力のうち第1の出力を入力し、第N+1のクロスバススイッチは、M個の入力端に、前記第1乃至第MのノードのN個の出力のうち第Nの出力を入力し、前記第I（ただし、Iは2以上且つN以下の整数）のクロスバススイッチは、M個の入力端に、前記第1乃至第Mのノードの出力のうち、前記第Iのクロスバススイッチに対応する順番の第I-1の出力と第Iの出力をそれぞれ入力とするM個の選択回路を備え、前記第Jのノード（ただし、Jは1以上且つM以下の整数）の入力端には、第1乃至第N+1のクロスバススイッチのうち隣り合う2つのクロスバ

スイッチの J 番目の出力をそれぞれ入力とする N 個の選択回路を備え、前記各選択回路は、クロスバスイッチの障害処理を行う障害処理回路から出力される選択制御信号により、2 入力的一方を選択出力し、一のクロスバスイッチに障害発生時、該障害クロスバスイッチを未使用とする。

【 0 0 1 1 】

本発明において、前記クロスバスイッチが、計算機内の CPU とメモリ間の接続、もしくは、CPU とメモリを備え前記クロスバスイッチを介してリモートノードのメモリへのアクセスを行うマルチノードシステムのノード間のスイッチングを行う。

【 0 0 1 2 】

【発明の実施の形態】

本発明の実施の形態について説明する。本発明は、システムに N 個必要なクロスバスイッチを、冗長分 1 個を含めて N + 1 個設け、システム障害発生時に障害処理回路がクロスバスイッチの障害を認識した場合、システム再立ち上げ後は、クロスバスイッチの入出力部分に設けた選択回路を、障害処理回路が制御することで、障害の発生したクロスバスイッチを未使用とし、冗長分のクロスバスイッチを使用するように、制御する。

【 0 0 1 3 】

より詳細には、本発明は、その好ましい一実施の形態において、第 1 乃至第 M のノード（図 1 では、M = 8 に対応）に対してノード間の接続に必要とされる N 個クロスバスイッチ（図 1 では、N = 8 に対応）にさらに 1 つのクロスバスイッチを備えた、第 1 乃至第 N + 1 のクロスバスイッチを備え、前記各ノードは、出力端から第 1 乃至第 N の出力信号を出力するとともに、入力端から、第 1 乃至第 N の入力信号を入力し、第 1 のクロスバスイッチ（図 1 のクロスバスイッチ 1 0）は前記第 1 乃至第 M のノードの各ノードの第 1 の出力信号を入力し、第 N + 1 のクロスバスイッチ（図 1 のクロスバスイッチ 1 8）は、前記第 1 乃至第 M のノードの各ノードの第 N の出力信号を入力し、第 I（ただし、I は 2 以上 N 以下）のクロスバスイッチ（図 1 のクロスバスイッチ 1 1 ~ 1 7）は、入力部に、前記第 1 乃至第 M のノードの各ノードの第 I - 1 の出力信号と第 I の出力信号の 2 つ

の信号をそれぞれ入力とする第1乃至第Mの選択回路（11-0～11-7、…、17-0～17-7）を備え、前記第J（ただし、Jは1以上M以下の整数）のノードのN個の入力端には、第1乃至第N+1のクロスバスイッチのうち隣り合うクロスバスイッチ、すなわち第Kと第K+1（ただし、Kは1以上でN以下の整数）のクロスバスイッチのJ番目の出力ポートの出力をそれぞれ入力とするM個の選択回路（0-0～0-7、1-0～1-7、…、7-0～7-7）を備え、これらの選択回路（11-0～11-7、…、17-0～17-7、0-0～0-7、1-0～1-7、…、7-0～7-7）は、クロスバスイッチの障害処理を行う障害処理回路（20）から出力される選択制御信号により、2つの信号の一方を選択出力することで、障害の発生したクロスバスイッチを未使用とし、冗長分のクロスバスイッチを使用するように制御する。

【0014】

本発明の一実施の形態において、前各ノードは、バイト単位に、第1乃至Nのデータを入出力する。なお、ビット単位であってもよい。

【0015】

障害処理回路は、N+1個のクロスバスイッチについて、システム制御部から出力された障害情報を保持するN+1ビットのクロスバ障害情報保持レジスタ（200）と、クロスバ障害情報保持レジスタ（200）の値に基づき、前記選択回路に対して選択制御信号を出力する選択回路制御出力回路（201）と、複数のクロスバスイッチの障害発生時、多重障害発生をシステム制御部に通知する多重障害検出部（202）と、を備えている。

【0016】

【実施例】

上記した本発明の実施の形態についてさらに詳細に説明すべく、本発明の実施例について図面を参照して以下に説明する。図1は、本発明の一実施例のシステム構成を示す図である。図1を参照すると、8個のノード0～7と、9個のクロスバスイッチ10～18と、障害処理回路20と、選択回路0-0～0-7、…、7-0～7-7、11-0～11-7、…、17-0～17-7を備えて構成されている。

【0017】

前記各ノード0～7は同一の構成とされ、また各クロスバスイッチ10～11は同一の構成とされている。各クロスバスイッチ10～18は、8個の入力ポートと、8個の出力ポートと、 8×8 のクロスバスイッチ部（不図示）と、クロスバスイッチ部に入力ポートと出力ポートの切り替えを制御する接続制御部（不図示）を備えており、各ポートは1バイト（8ビット）単位の入出力を行う構成とされる。なお、図1は、あくまで、本発明を説明するためのものであり、本発明において、例えばノードは8個に限定されるものでないことは勿論である。

【0018】

ノード0～7において、任意の2つのノード間の通信データは、クロスバスイッチ10～18により、通信元ノードから通信先ノードへ転送される。

【0019】

2つのノード間のデータ通信のデータ幅は例えば8バイト（ $8 \times 8 = 64$ ビット）である。

【0020】

クロスバスイッチ10には、各ノード0～7からそれぞれ出力される8バイトデータにおけるバイト0データが8本の入力ポートにそれぞれ入力される。

【0021】

クロスバスイッチ11においては、ノード0から出力されるバイト0データとバイト1データが選択回路11-0に入力され選択回路11-0の出力が、クロスバスイッチ11の第1入力ポートに入力され、ノード1から出力されるバイト0データとバイト1データが選択回路11-1に入力され選択回路11-1の出力が、クロスバスイッチ11の第2入力ポートに入力され、以下同様にして、ノード7から出力されるバイト0データとバイト1データが選択回路11-7に入力され選択回路11-7の出力が、クロスバスイッチ11の第8入力ポートに入力される。

【0022】

選択回路11-0～11-7は、障害処理回路20からの制御信号により、ノ

ード0～7から出力された8バイトデータにおけるバイト0データとバイト1データの一方を選択し、クロスバスイッチ11へ出力する。選択回路11-0～11-7は、例えば障害無しの場合は、バイト1データを選択し、クロスバスイッチ10の障害時には、バイト0データを選択する（後述する図4参照）。

【0023】

クロスバスイッチ17においては、ノード0から出力されるバイト6データとバイト7データが選択回路17-0に入力され選択回路17-0の出力が、クロスバスイッチ17の第1の入力ポートに入力され、ノード1から出力されるバイト6データとバイト7データが選択回路17-1に入力され選択回路17-1の出力が、クロスバスイッチ17の第2の入力ポートに入力され、以下同様にして、ノード7から出力されるバイト6データとバイト7データが選択回路17-7に入力され選択回路17-7の出力が、クロスバスイッチ17の第8の入力ポートに入力される。

【0024】

クロスバスイッチ18には、各ノード0～7からそれぞれ出力される8バイトデータにおけるバイト7データが8本の入力ポートにそれぞれ入力される。

【0025】

クロスバスイッチ10～18から出力されたデータは、それぞれ選択回路0-0～0-7、1-0～1-7、…、7-0～7-7により選択され、ノード0～7へ入力される。

【0026】

ノード0に対応する選択回路0-0は、クロスバスイッチ10の第1出力ポートとクロスバスイッチ11の第1出力ポートから出力されたバイト0データを入力し、障害処理回路20からの制御信号に基づき一方を選択してノード0へ出力する。選択回路0-7はクロスバスイッチ17の第1出力ポートとクロスバスイッチ18の第1出力ポートから出力されたバイト7データを入力し、障害処理回路20からの制御信号に基づき一方を選択してノード0へ出力する。

【0027】

同様にして、ノード7に対応する選択回路7-0は、クロスバスイッチ10の

第8出力ポートとクロスバススイッチ11の第8出力ポートから出力されたバイト0データを入力し、障害処理回路20からの制御信号に基づき一方を選択してノード0へ出力する。選択回路7-7は、障害処理回路20からの制御信号により、クロスバススイッチ17の第8出力ポートとクロスバススイッチの第8出力ポート18から出力されたバイト7データを選択し、ノード7へ出力する。

【0028】

障害処理回路20は、発生した障害に関する障害情報に基づき、選択回路0-0~0-7、1-0~1-7、…、7-0~7-7、11-0~11-7、…、17-0~17-7へ選択制御信号を出力する。

【0029】

図2は、図1におけるノード0~ノード7の内部構成の一例を示したものである。各ノードは、4個のCPU100~103と、メモリコントローラ104と、メモリ105と、I/Oコントローラ106とを備えて構成されている。

【0030】

各CPU100~103は、メモリコントローラ104を経由してメモリアクセス、I/Oアクセスを行う。

【0031】

自ノード内メモリ105へのアクセスの場合には、メモリコントローラ104からメモリ105へアクセスが行われるが、他ノード内メモリへのアクセスの場合には、メモリコントローラ104からクロスバススイッチを経由して、リクエストが他ノードのメモリコントローラへ送られ、他ノード内メモリへのアクセスが行われる。

【0032】

図3は、図1における障害処理回路20の内部構成を示したものである。障害処理回路20は、例えばシステムの再立ち上げ後に、選択回路0-0~7-7、11-0~17-7に対して、選択制御信号を出力することで、障害の発生したクロスバススイッチを未使用とし、冗長分のクロスバススイッチを使用するように制御する。

【0033】

システム制御部から出力されたクロスバスイッチ10～18に関する障害情報は、9ビットのクロスバ障害情報保持レジスタ200へ入力され、各ビットはそれぞれクロスバスイッチ10～18の障害の有無を保持する。

【0034】

クロスバ障害情報保持レジスタ200の情報は、選択回路制御出力回路201へ出力され、選択回路制御出力回路201は、該情報に基づき、選択回路0-0～7-7、11-0～17-7に対して、それぞれ選択制御信号を出力する。

【0035】

クロスバ障害情報保持レジスタ200の情報は、クロスバ多重障害検出回路202へも出力され、クロスバスイッチ10～18のうちの2個以上に障害が発生した場合、クロスバ多重障害検出回路202は、多重障害発生をシステム制御部へ通報する。

【0036】

図4は、本実施例において、クロスバスイッチ10～18に障害が発生した時に、ノード間転送データの各バイトが、どのクロスバスイッチでスイッチング制御されるか、その一覧を示したものである。

【0037】

通常の障害のない状態では、図4の最下行に示されているように、バイト0～バイト7のデータは、それぞれ、クロスバスイッチ10～17により、スイッチングされる。

【0038】

そして、例えばクロスバスイッチ10に障害が発生した場合は、図4の2行目に示されているように、バイト0～バイト7のデータは、それぞれクロスバスイッチ11～18で、スイッチングされるように制御される。

【0039】

クロスバスイッチ11～18に障害が発生した場合も同様に、障害の発生したクロスバスイッチを回避して、各バイトデータが、それぞれ、図4に示されるクロスバスイッチによりスイッチングされるように制御される。

【0040】

次に、本発明の一実施例の動作について説明する。

【0041】

図1を参照すると、クロスバスイッチ10～18は、ノード間通信のための冗長構成のクロスバスイッチであり、障害が発生していない場合、クロスバスイッチ10～17が使用され、クロスバスイッチ18は使用されない。

【0042】

通常時は、ノード0～7から出力された8バイトデータにおけるバイト0データは、クロスバスイッチ10で、バイト1データはクロスバスイッチ11で、バイト7データはクロスバスイッチ17でデータのスイッチングが行われる。

【0043】

ノード0内のCPU100が、リモートノードであるノード1内のメモリへアクセスを行う場合、8バイトのリクエストデータにおけるバイト0データは、クロスバスイッチ10によりスイッチングされてノード1へ送付される。

【0044】

バイト0データは、ノード0から、選択回路11-0へも送付されるが、選択回路11-0は、障害処理回路20からの選択制御信号により、他方の入力であるノード0からの8バイトデータにおけるバイト1データを選択して出力する。

【0045】

クロスバスイッチ10から出力されたバイト0データは、選択回路1-0に入力され、選択回路1-0は、障害処理回路20からの選択制御信号により、バイト0データを選択し、ノード1へ出力する。

【0046】

システムに障害が発生し、障害発生後の診断処理の結果、クロスバスイッチ10に障害があることが判明した場合、システム再立ち上げの後、障害処理回路20からの選択回路へ出力される選択制御信号により、クロスバスイッチ10を未使用状態とし、ノード0～7から出力された8バイトデータにおける、バイト0データは、クロスバスイッチ11で、バイト1データは、クロスバスイッチ12で、バイト7データは、クロスバスイッチ18でデータのスイッチングが行われる。

【 0 0 4 7 】

この場合、ノード 0 からノード 1 へのデータ転送は、ノード 0 から、選択回路 1 1 - 0 へ出力されたバイト 0 データが、障害処理回路 2 0 からの選択制御信号により選択され、クロスバスイッチ 1 1 へ出力される。

【 0 0 4 8 】

クロスバスイッチ 1 1 から出力されたバイト 0 データは、選択回路 1 - 0 に入力され、選択回路 1 - 0 は、障害処理回路 2 0 からの選択制御信号により、バイト 0 データを選択し、ノード 1 へ入力される。

【 0 0 4 9 】

クロスバスイッチ 1 1 ~ 1 8 のいずれか一のスイッチに障害が発生した場合は、ノード転送データの各バイトは、図 4 で示されるクロスバスイッチを経由して、上記と同様な制御によりデータ転送が行われる。

【 0 0 5 0 】

クロスバスイッチ 1 0 ~ 1 8 のうちの 2 個以上で障害が発生した場合は、障害処理回路 2 0 におけるクロスバ多重障害検出回路 2 0 2 がクロスバ多重障害を検出し、システム制御部へ多重障害を通報する。この場合はシステムの再立ち上げが行われず、修理が行われるまでシステムダウンとなる。

【 0 0 5 1 】

なお、上記実施例では、各ノードは 8 バイトデータを出力する構成とし、各選択回路、クロスバスイッチの各ポートは、1 バイト単位でデータの入出力を行う構成を示したが、本発明は、かかる構成に限定されるものではなく、ワード単位、あるいはビット単位でデータの入出力を行う構成としてもよいことは勿論である。

【 0 0 5 2 】

また本発明は、マルチノード構成の計算機システムに適用して好適とされるほか、複数の CPU とメモリ間の接続を制御するクロスバスイッチに対しても、同様に適用できる。

【 0 0 5 3 】

【発明の効果】

以上説明したように、本発明によれば下記記載の効果を奏する。

【 0 0 5 4 】

本発明の第 1 の効果は、クロスバススイッチを冗長構成とし、クロスバススイッチに障害が発生した場合は、障害情報に基づき、障害処理回路が各クロスバススイッチの入力部、および出力部に設けられた選択回路を制御することで、システム再立ち上げ後、障害のあるクロスバススイッチを回避した運用を行うことを可能とする、ということである。

【 0 0 5 5 】

本発明の第 2 の効果は、クロスバススイッチの修理を実施するまでシステムの復旧ができないという状態を回避することが可能となり、前記状態を回避するために、クロスバススイッチを完全二重化する必要もない、ということである。

【 0 0 5 6 】

本発明の第 3 の効果として、クロスバススイッチを活線挿抜仕様にした場合は、オンライン状態で、クロスバススイッチの部品交換を行うことが可能であり、システムを止めずに保守作業を可能とする、ということである。

【 0 0 5 7 】

本発明の第 4 の効果として、クロスバススイッチ障害時の切り替えに際して、データの分岐入力に隣り合うクロスバススイッチ間で行われるため、上記特開平 1 1 - 3 3 1 3 7 4 号公報等のシステムで問題となった、迂回によるデータ遅延時間の変動等は生じないか、無視できる程度のものであり、高速動作周波数の計算機システムに適用可能である、ということである。

【図面の簡単な説明】

【図 1】

本発明の一実施例のシステム構成を示す図である。

【図 2】

本発明の一実施例の CPU の構成を示す図である。

【図 3】

本発明の一実施例の障害処理回路の構成を示す図である。

【図 4】

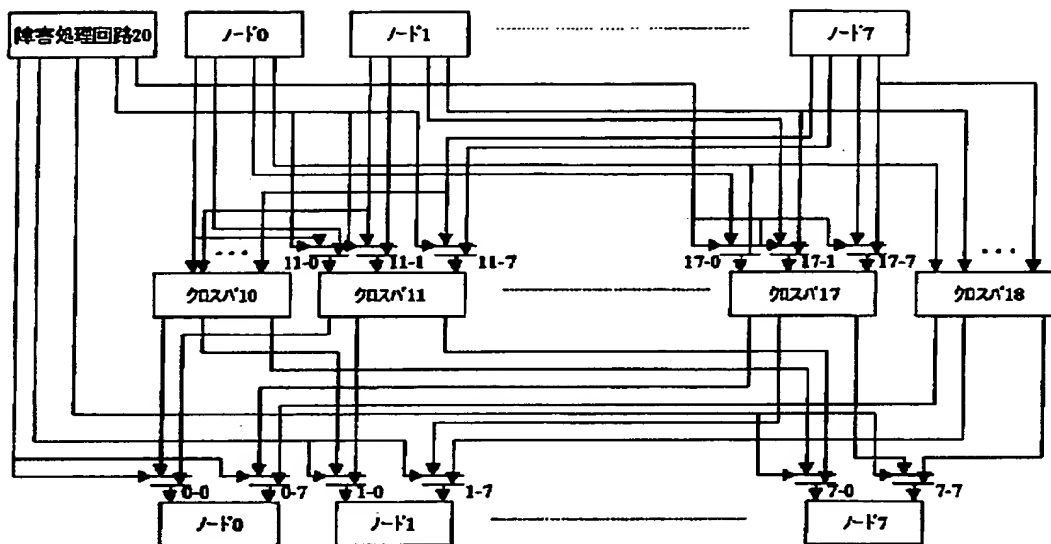
本発明の一実施例において、障害クロスバスイッチと、転送データの各バイトが経由するクロスバスイッチの関係を一覧で示す図である。

【符号の説明】

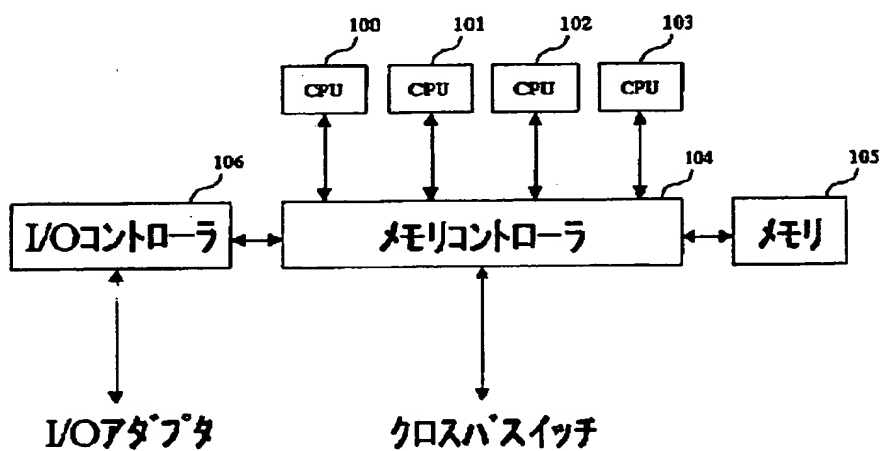
- 0 ～ 7 ノード
- 0 - 0 ～ 7 - 7 選択回路
- 1 0 ～ 1 8 クロスバスイッチ
- 2 0 障害処理回路
- 1 0 0 ～ 1 0 3 C P U
- 1 0 4 メモリコントローラ
- 1 0 5 メモリ
- 1 0 6 I / O コントローラ
- 2 0 0 クロスバ障害情報保持レジスタ
- 2 0 1 選択回路制御出力回路
- 2 0 2 クロスバ多重障害検出回路

【書類名】 図面

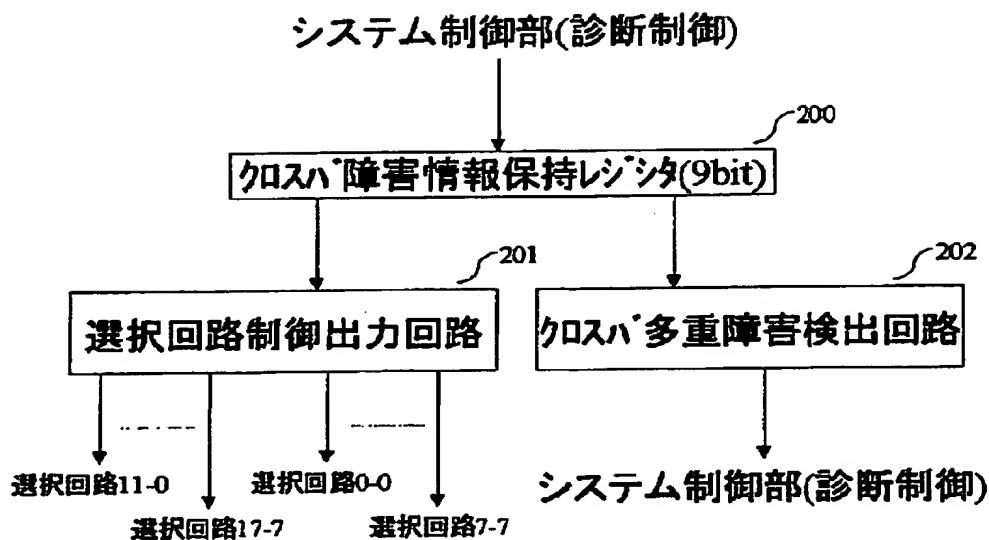
【図 1】



【図 2】



【図 3】



【図 4】

障害クロスバ	バイト0	バイト1	バイト2	バイト3	バイト4	バイト5	バイト6	バイト7
クロスバ10	クロスバ11	クロスバ12	クロスバ13	クロスバ14	クロスバ15	クロスバ16	クロスバ17	クロスバ18
クロスバ11	クロスバ10	クロスバ12	クロスバ13	クロスバ14	クロスバ15	クロスバ16	クロスバ17	クロスバ18
クロスバ12	クロスバ10	クロスバ11	クロスバ13	クロスバ14	クロスバ15	クロスバ16	クロスバ17	クロスバ18
クロスバ13	クロスバ10	クロスバ11	クロスバ12	クロスバ14	クロスバ15	クロスバ16	クロスバ17	クロスバ18
クロスバ14	クロスバ10	クロスバ11	クロスバ12	クロスバ13	クロスバ15	クロスバ16	クロスバ17	クロスバ18
クロスバ15	クロスバ10	クロスバ11	クロスバ12	クロスバ13	クロスバ14	クロスバ16	クロスバ17	クロスバ18
クロスバ16	クロスバ10	クロスバ11	クロスバ12	クロスバ13	クロスバ14	クロスバ15	クロスバ17	クロスバ18
クロスバ17	クロスバ10	クロスバ11	クロスバ12	クロスバ13	クロスバ14	クロスバ15	クロスバ16	クロスバ18
クロスバ18	クロスバ10	クロスバ11	クロスバ12	クロスバ13	クロスバ14	クロスバ15	クロスバ16	クロスバ17
障害無し	クロスバ10	クロスバ11	クロスバ12	クロスバ13	クロスバ14	クロスバ15	クロスバ16	クロスバ17

【書類名】 要約書

【要約】

【課題】

クロスバスイッチに障害が発生した場合の早急なシステムの復旧を、比較的安価なシステムで実現するシステムの提供。

【解決手段】

N + 1 の冗長構成のクロスバスイッチを備え、第 1 のクロスバスイッチは N 個の入力端に、各ノードの第 1 の出力、第 N + 1 のクロスバスイッチは各ノードの第 N の出力を入力し、第 2 ～第 N のクロスバスイッチ（I 番目）の入力端には、各ノードの出力のうち I - 1 番目と I 番目の隣り合う出力を入力とする第 1 乃至第 N の選択回路を備え、各ノード（J 番目）の入力端には第 1 ～第 N + 1 のクロスバスイッチの隣り合う二つのクロスバスイッチの J 番目の出力をそれぞれ入力とする N 個の選択回路を備え、各選択回路はクロスバスイッチの障害処理を行う障害処理回路から出力される選択制御信号により 2 入力的一方を選択出力する。

【選択図】

図 1

出 願 人 履 歴 情 報

識別番号 [0 0 0 0 0 4 2 3 7]

1. 変更年月日 1 9 9 0 年 8 月 2 9 日

[変更理由] 新規登録

住 所 東京都港区芝五丁目 7 番 1 号

氏 名 日本電気株式会社